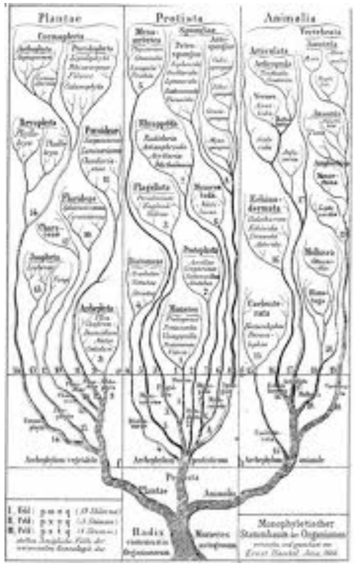
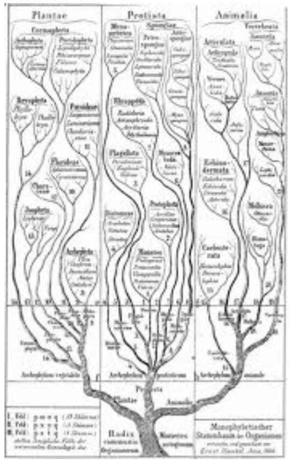
Long-term diversity dynamics: reconciling molecular phylogenies with the fossil record



Hélène Morlon

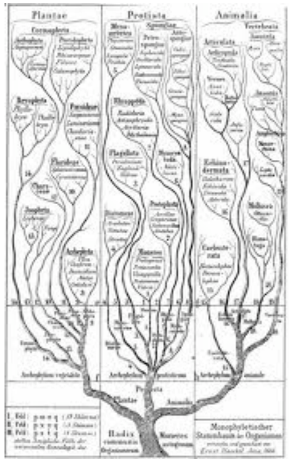
CNRS, Ecole Polytechnique

Outline



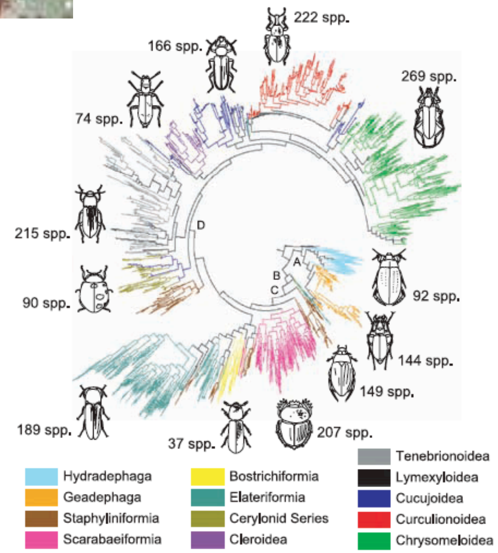
- My interest in studying long-term diversity dynamics
- How to study long-term diversity dynamics?
- Current phylogenetic inference is typically inconsistent with the fossil record
- Reconciling molecular phylogenies with the fossil record
- Conclusions
- Current related project & future directions

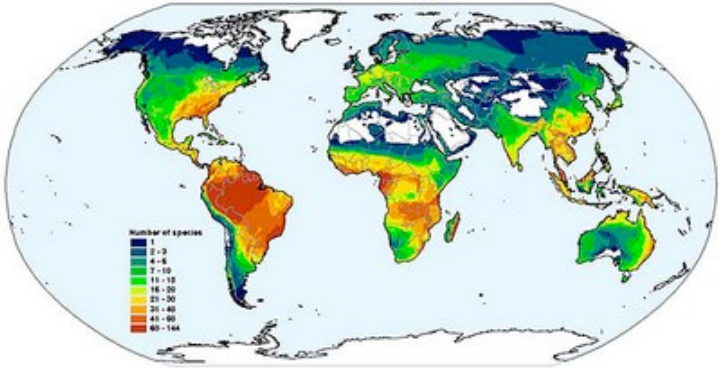
Outline



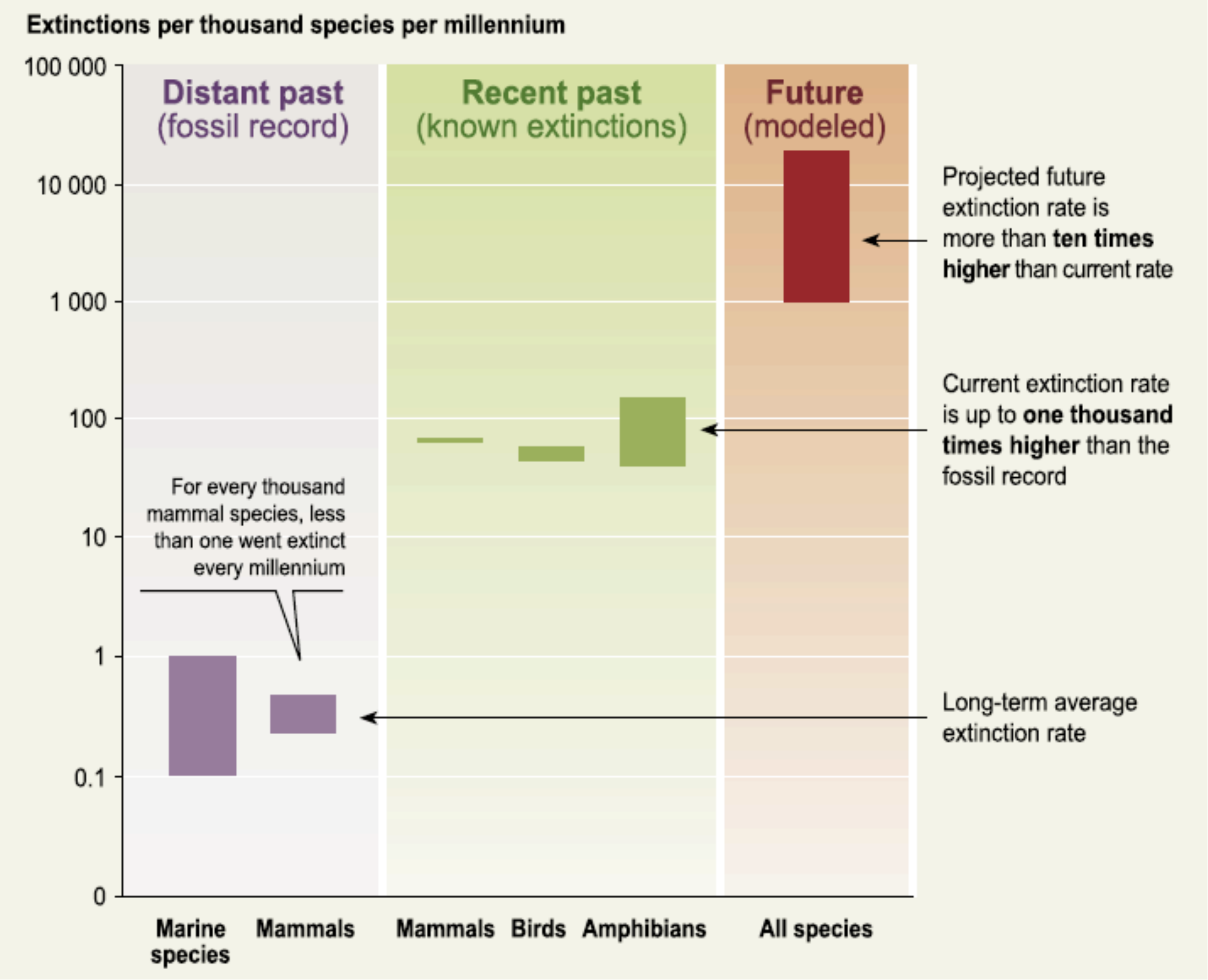
- **My interest in studying long-term diversity dynamics**
- How to study long-term diversity dynamics?
- Current phylogenetic inference is typically inconsistent with the fossil record
- Reconciling molecular phylogenies with the fossil record
- Conclusions
- Current related project & future directions

What are the long-term diversity dynamics driving present-day biodiversity patterns?



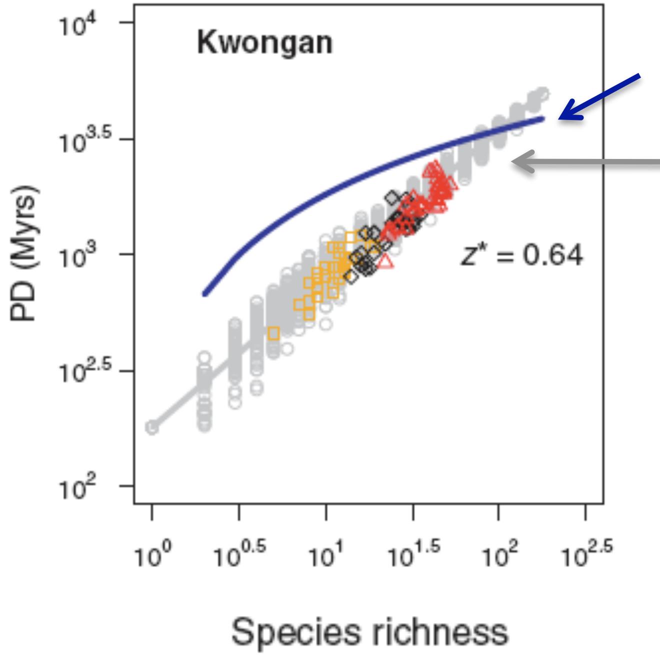


Estimating background extinction rates



Millenium Ecosystem Assessment

Estimating the loss of evolutionary history with extinctions or habitat loss



diversity at equilibrium

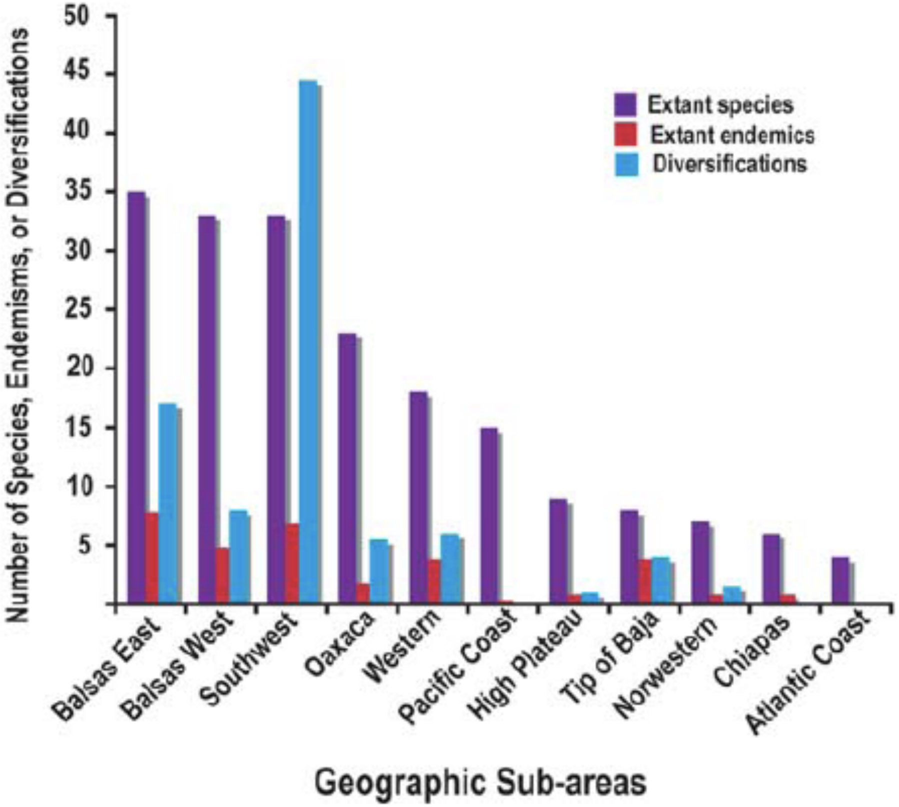
diversity in expansion

Morlon *et al*. *Ecol Lett* 2011

Long-term diversity dynamics influence tree shape and thus the amount of phylogenetic diversity lost by extinction

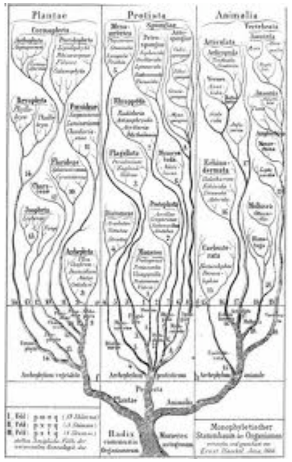
Nee & May *Science* 1997

Identifying source and sink areas for diversification



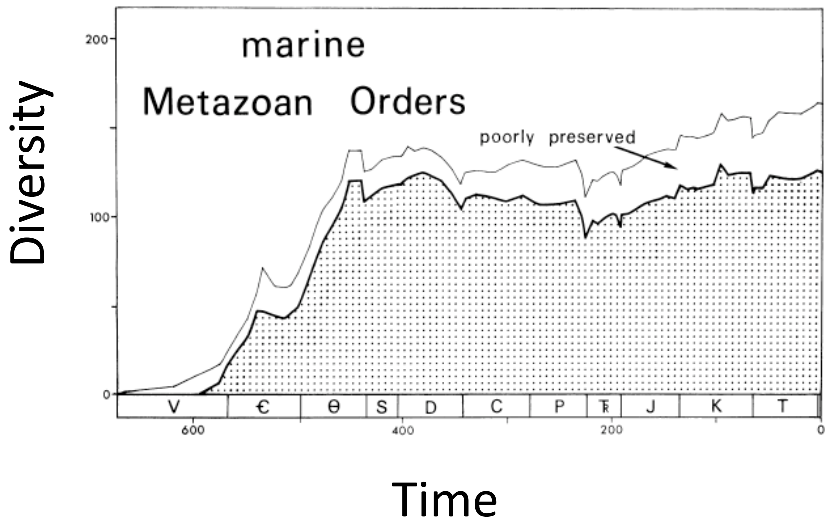
Beccera & Venable
***PloS One* 2008**

Outline

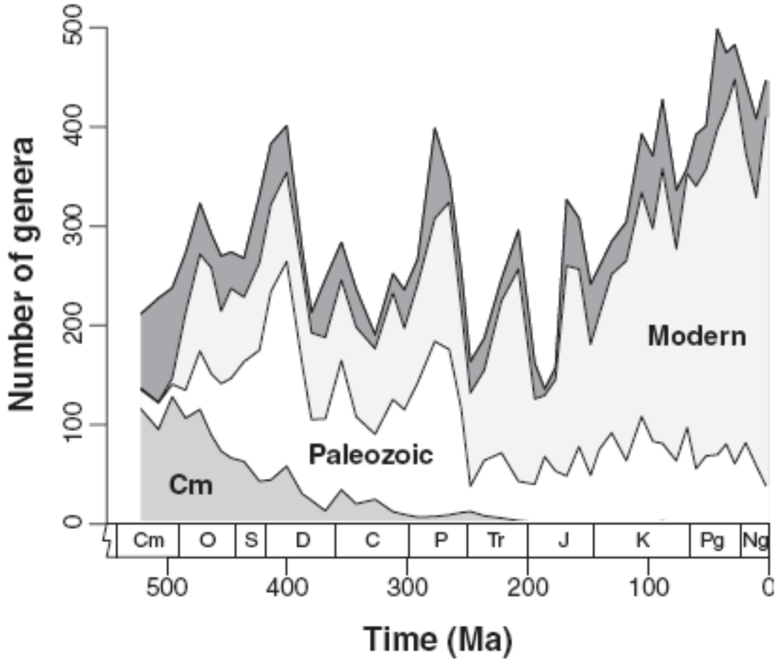


- My interest in studying long-term diversity dynamics
- **How to study long-term diversity dynamics?**
- Current phylogenetic inference is typically inconsistent with the fossil record
- Reconciling molecular phylogenies with the fossil record
- Conclusions
- Current related project & future directions

Diversity dynamics may be inferred using fossil data



Sepkoski *Paleobiology* 1978



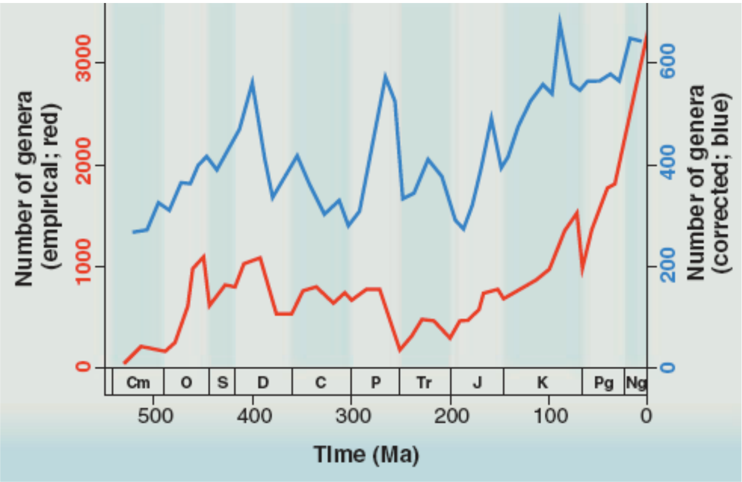
Alroy *Science* 2010

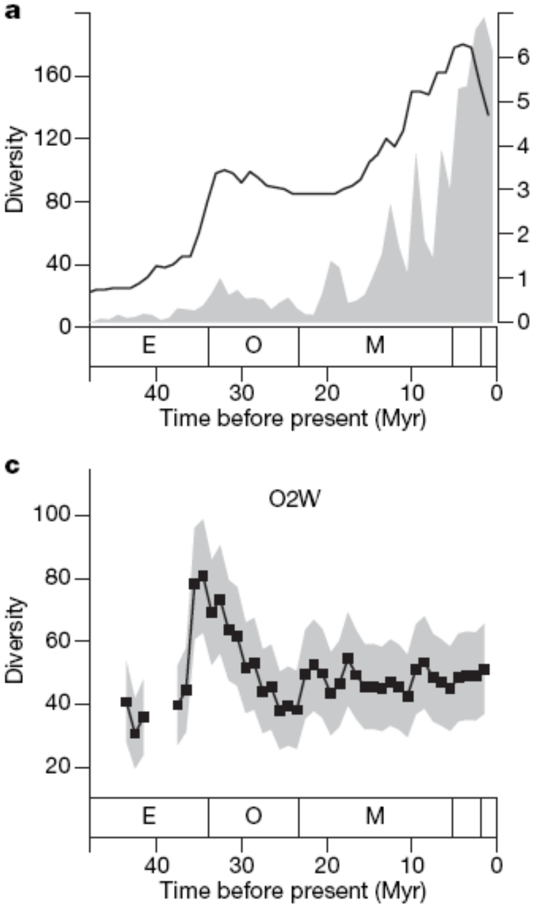
Many groups lack fossil data

Different methods to correct for the incompleteness of the fossil record yield contrasting results

Planktonic diatoms

Marine invertebrates





Benton *Science* 2009

Rabosky & Sorhannus
Nature 2009

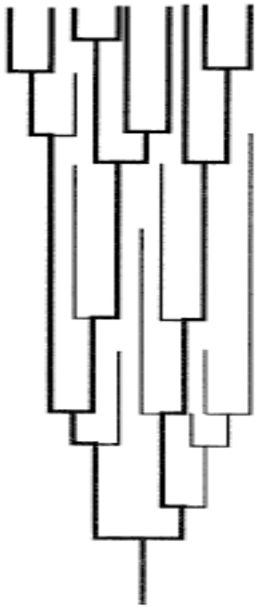
Analyzing diversity dynamics using the molecular phylogenies of extant species:

the bases

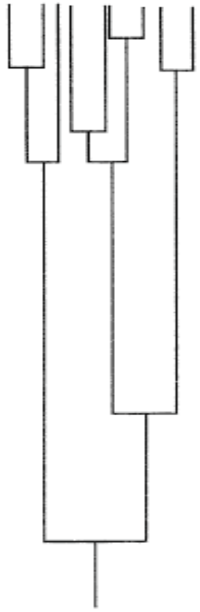
Sean Nee

Paul Harvey

actual lineages



apparent lineages



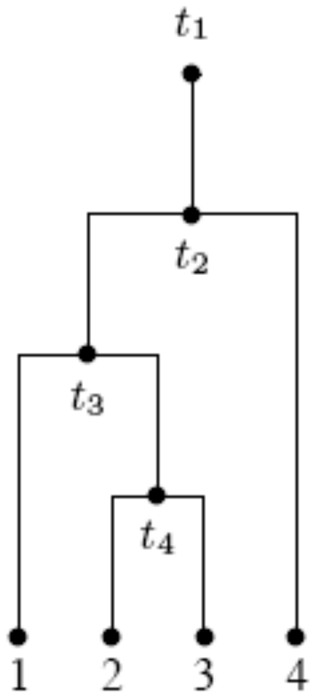
The reconstructed evolutionary process

SEAN NEE, ROBERT M. MAY AND PAUL H. HARVEY

PTB 1994

Analyzing diversity dynamics using the molecular phylogenies of extant species:

the bases



We assume that at time t_1 in the past a single lineage starts to diversify according to a birth-death process: each lineage gives rise to a new lineage with rate λ, and go extinct with rate μ (λ and μ are constant in time and across lineages)

Then it is possible to derive the likelihood of observing the splitting times $\{t_i\}$ in the reconstructed phylogeny, and thus to estimate λ and μ by maximum likelihood

Molecular phylogenies are widely used to infer diversity dynamics

Review

TRENDS in Ecology and Evolution Vol.22 No.11

ELSEVIER

Full text provided by www.sciencedirect.com

ScienceDirect

Estimating diversification rates from phylogenetic information

Robert E. Ricklefs

Department of Biology, University of Missouri-St Louis, MO 63121-4499, USA

TREE 2007

When λ and μ vary over time

EXPLOSIVE EVOLUTIONARY RADIATIONS: DECREASING SPECIATION OR INCREASING EXTINCTION THROUGH TIME?

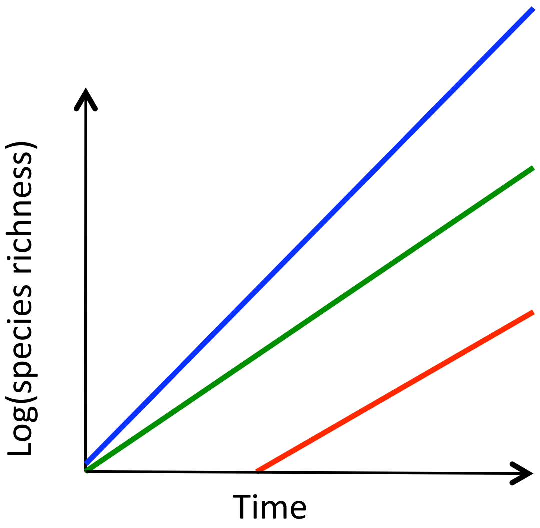
Daniel L. Rabosky[1,2,3] and Irby J. Lovette[1,3]

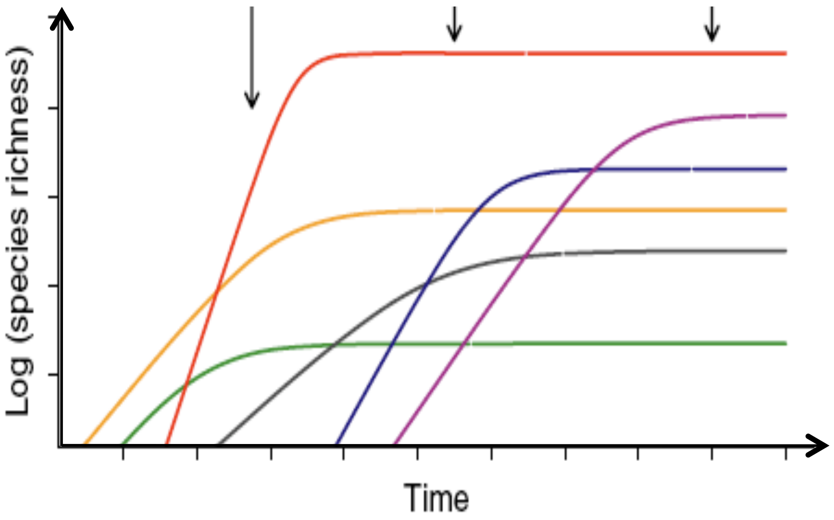
Evolution 2008

Some limitations

What if phylogenies are incomplete?

How to test the hypothesis that diversity is at equilibrium?
(we have assumed that diversity is expanding)





OPEN ACCESS Freely available online

PLoS BIOLOGY

Inferring the Dynamics of Diversification: A Coalescent Approach

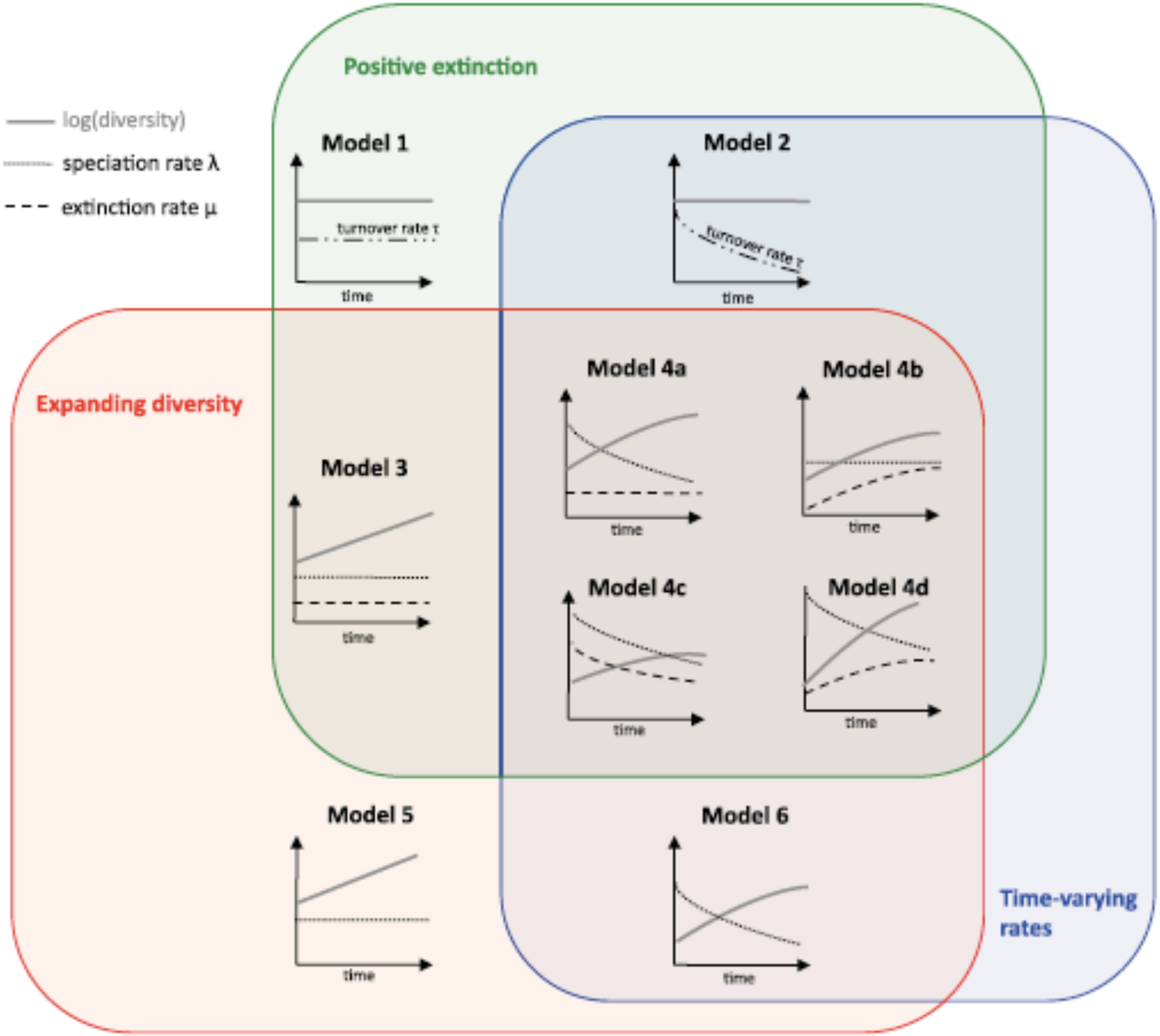
Hélène Morlon[1]*, Matthew D. Potts[2], Joshua B. Plotkin[1]*

PloSB 2010

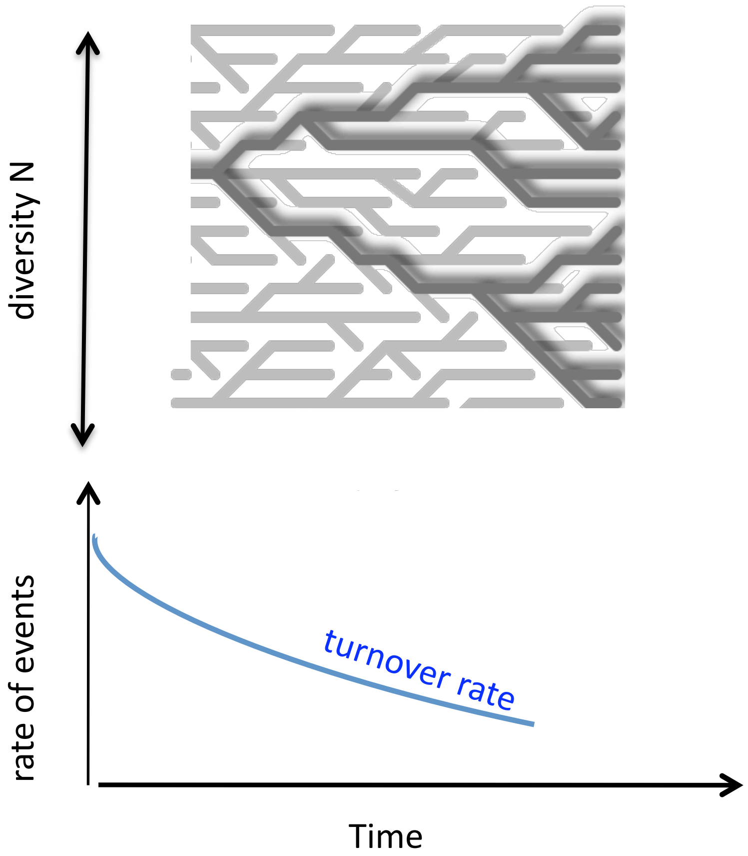
1. *Accomodating incomplete phylogenies*

2. *testing the hypothesis that diversity is at equilibrium*
   (while accomodating incomplete phylogenies
   and rate variation over time)

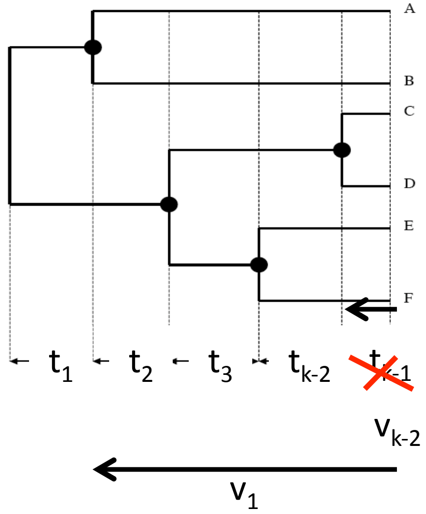
9 diversification scenarios



Likelihood of a phylogeny under a given diversification scenario Case of constant diversity



randomly sample k species out of N



$$\Lambda(t_1,t_2,\ldots,t_{k-2}) = \sum_{i=1}^{k-2} \log L(t_i)$$

turnover rate

$$L(t_i) = \frac{i(i+1)}{2}\frac{2\lambda(v_i)}{N}\exp\left[-\frac{i(i+1)}{2}\int_{v_i-t_i}^{v_i}\frac{2\lambda(t)}{N}dt\right]$$

Morlon *et al.* *PloSB* 2010

Griffiths & Tavare PTB 1997

Likelihood of a phylogeny under a given diversification scenario **Case of varying diversity**

1. Assume that a stochastic birth-death process underlies cladogenesis: speciation rate λ, extinction rate μ, potentially varying through time

2. Consider the expected variation in diversity through time N(t), given present-day diversity (time is measured from the present to the past)

Then the likelihood of internode distances for the phylogeny of k randomly sampled species is given by:

$$\Lambda(t_1,t_2,\ldots,t_{k-1}) = \sum_{i=1}^{k-1} \log L(t_i)$$

speciation rate

$$L(t_i) = \frac{i(i+1)}{2}\,\frac{2\lambda(v_i)}{N(v_i)}\exp\left[-\frac{i(i+1)}{2}\int_{v_i-t_i}^{v_i}\frac{2\lambda(t)}{N(t)}\,dt\right]$$

Morlon *et al. PloSB* 2010

Likelihood of a phylogeny under a given diversification scenario

Case of varying diversity

1. Assume that a stochastic birth-death process underlies cladogenesis: speciation rate λ, extinction rate μ, potentially varying through time

2. Consider the expected variation in diversity through time N(t), given present-day diversity (time is measured from the present to the past)

Then the likelihood of internode distances for the phylogeny of k randomly sampled species is given by:

$$\Lambda(t_1, t_2, \ldots, t_{k-1}) = \sum_{i=1}^{k-1} \log L(t_i)$$

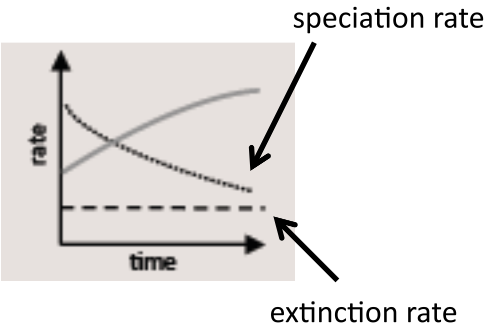
N is stochastic,
but approximated by
its deterministic expectation

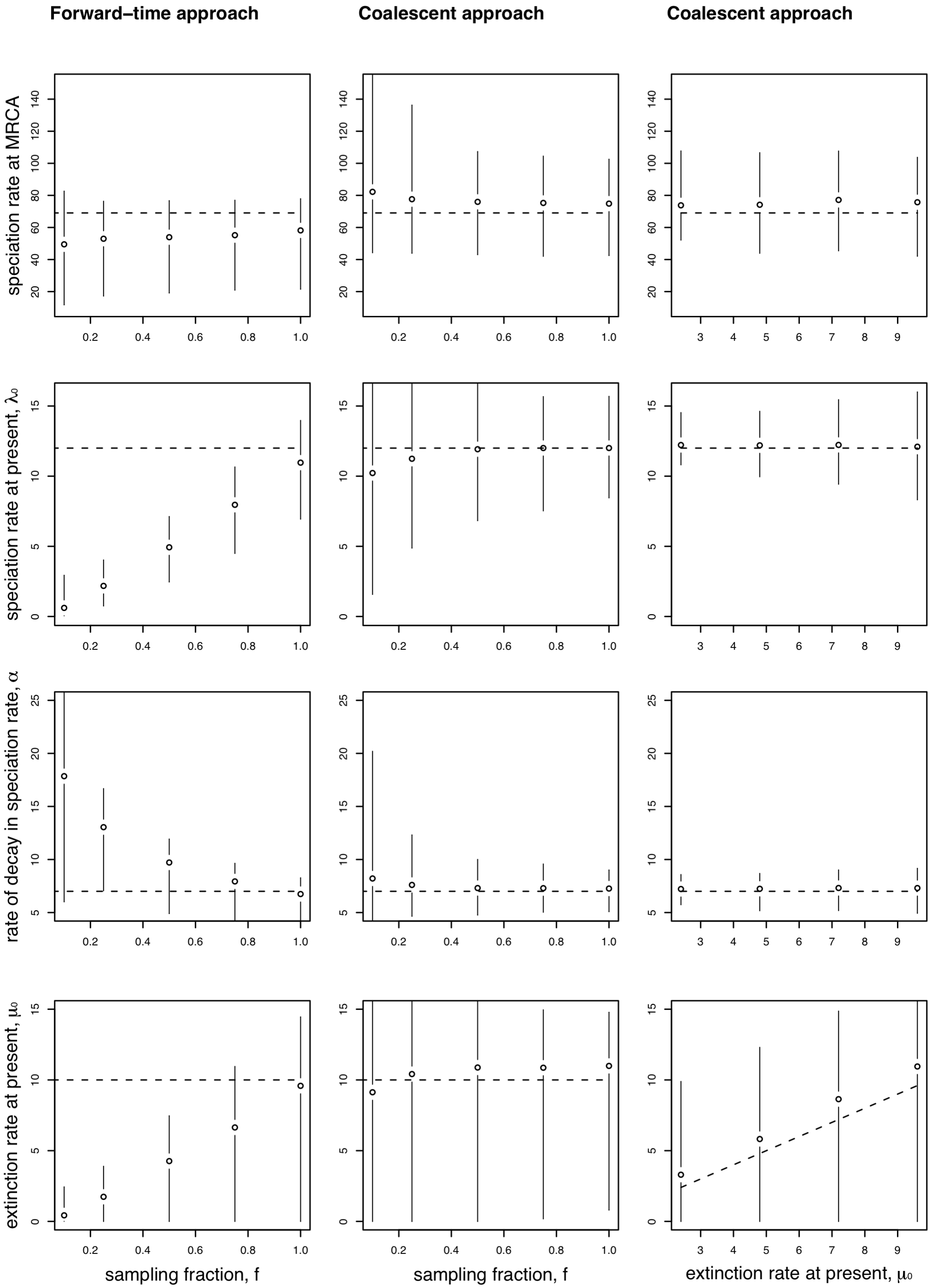
speciation rate

$$L(t_i) = \frac{i(i+1)}{2} \frac{2\lambda(v_i)}{N(v_i)} \exp\left[-\frac{i(i+1)}{2} \int_{v_i - t_i}^{v_i} \frac{2\lambda(t)}{N(t)} dt\right]$$

Morlon *et al. PloSB* 2010

The coalescent provides a robust approach for estimating diversification rates, even when phylogenies are only partially sampled





Morlon *et al.* *PloSB* 2010

Application : meta-analysis on 289 phylogenies

McPeek's repository (*Am. Nat.* 2007 & 2008)

Phylogenies from the recent literature.
Include chordate, arthropod, mollusk, and magnoliophyte phylogenies
Minimum size: 4 species; maximal size: 116 species
Min fraction sampled: 36%; max fraction sampled: 100%

Phillimore's bird phylogenies (*PloSB* 2008)

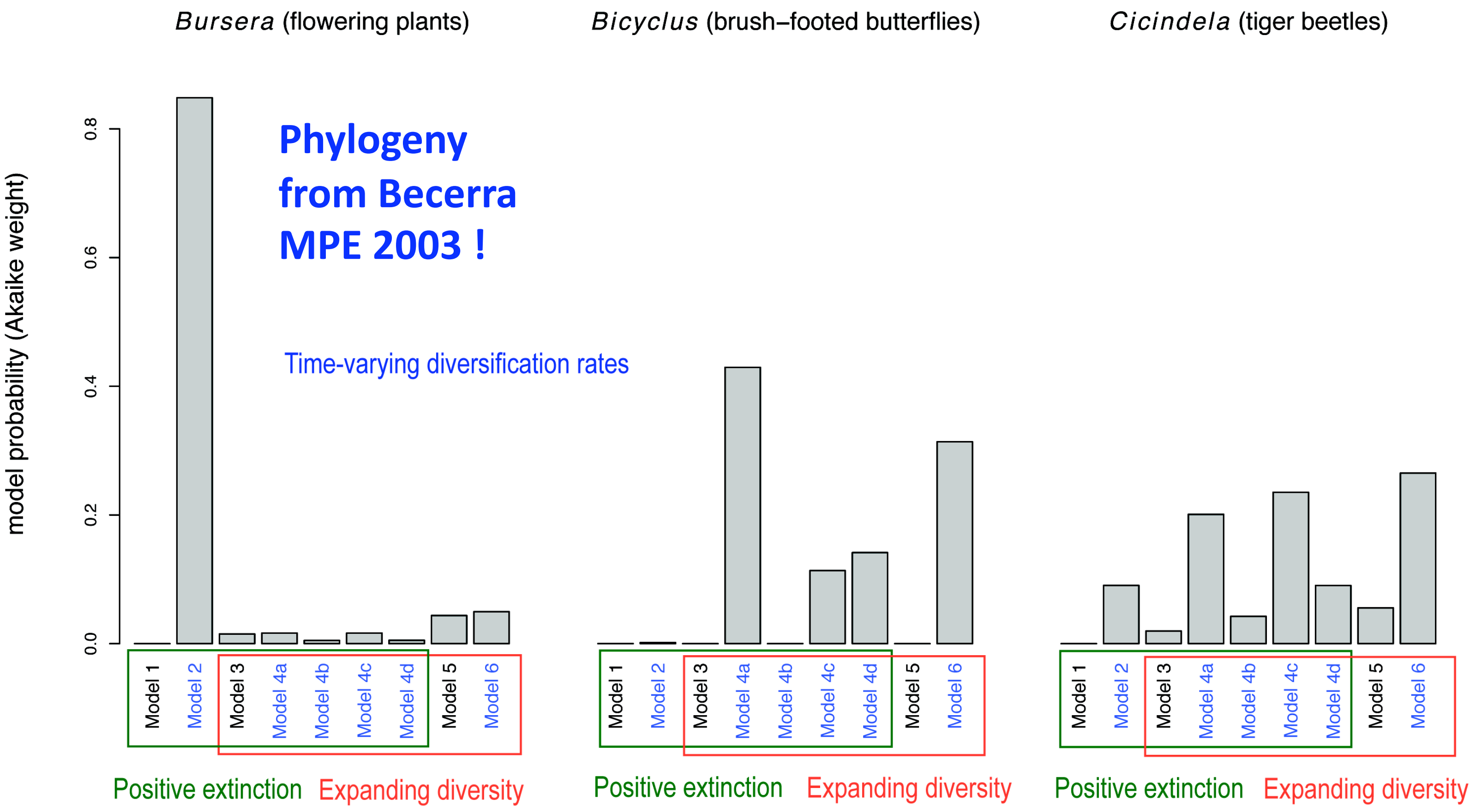
Molecular phylogenies, constructed using *BEAST*
Minimum size: 7 species; maximal size: 59 species
Min fraction sampled: 50%; max fraction sampled: 100%

Morlon *et al. PloSB* 2010

Dynamics of diversification in example phylogenies

$$\text{Akaike weight} = \frac{\exp(-\Delta_i/2)}{\sum_{r=1}^{R}\exp(-\Delta_r/2)}$$

« probability that a model is the best among the whole set of candidate models »

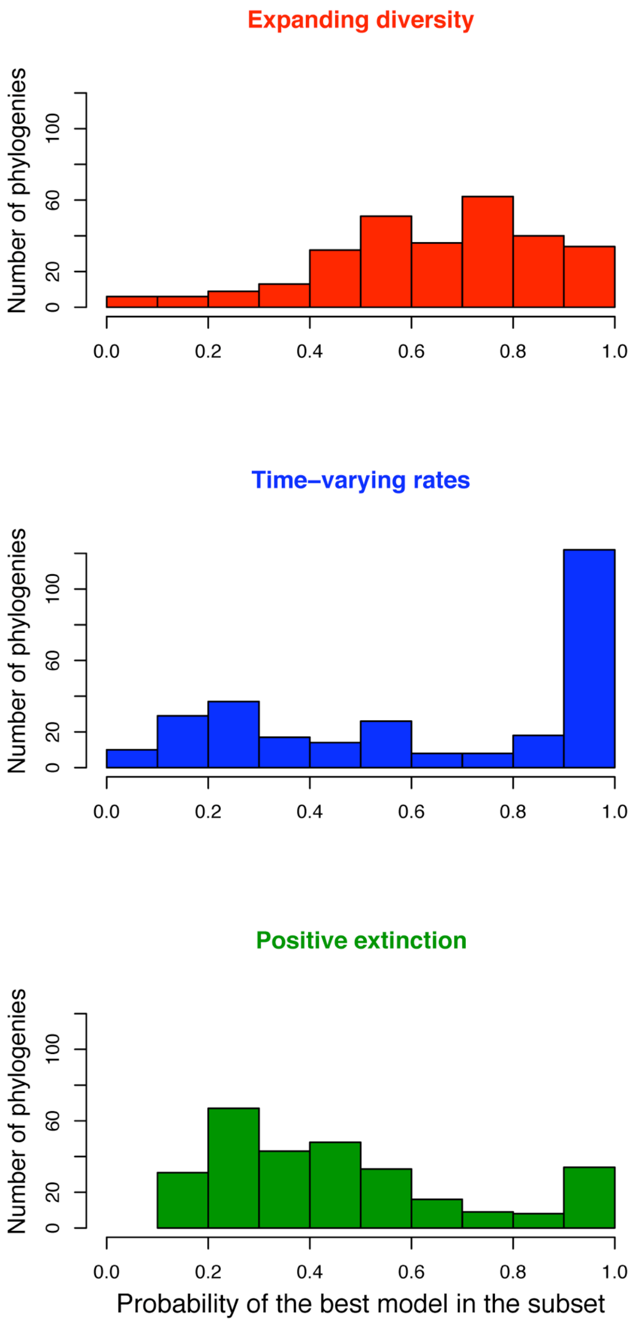


Phylogeny from Becerra MPE 2003 !

Morlon *et al. PloSB* 2010

Most phylogenies are consistent with the hypothesis that diversity is expanding with time- varying diversification rates

Suggest the preponderance of adaptive radiations and ecological limits on diversification



Morlon *et al. PloSB* 2010

Density-dependent cladogenesis in a bacterial species group

Dustin Brisson
UPenn

Rate variation through time, which has been interpreted as a signal of density-dependent cladogenesis and adaptive radiation, has been widely reported for macroorganisms. What about microorganisms?

Morlon *et al. in prep.*

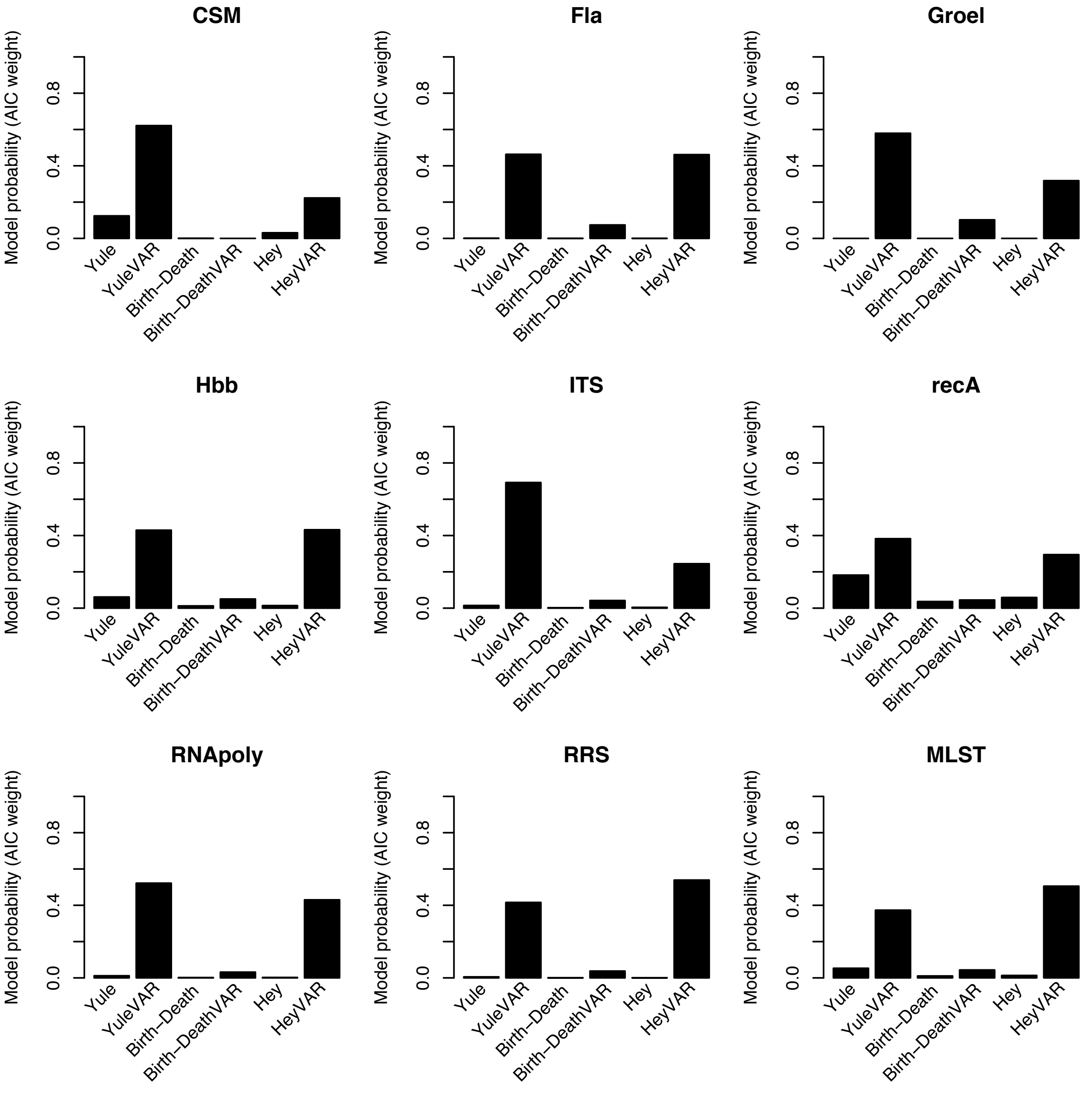
Density-dependent cladogenesis in a bacterial species group

Borrelia burgdorferi: obligate parasite with ~15 recognized phylotypes, vectored among vertebrate hosts by ticks. 3 of them can cause human Lyme disease

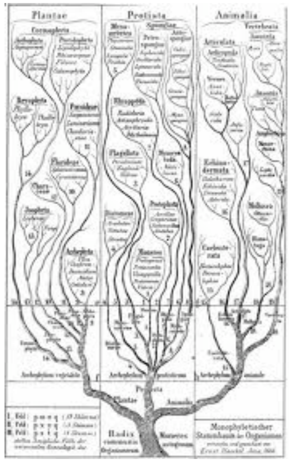
Phylogenies: constructed from partial genome sequence data, 6 different loci, and the loci combined

Morlon *et al. In prep.*

Density-dependent cladogenesis in a bacterial species group

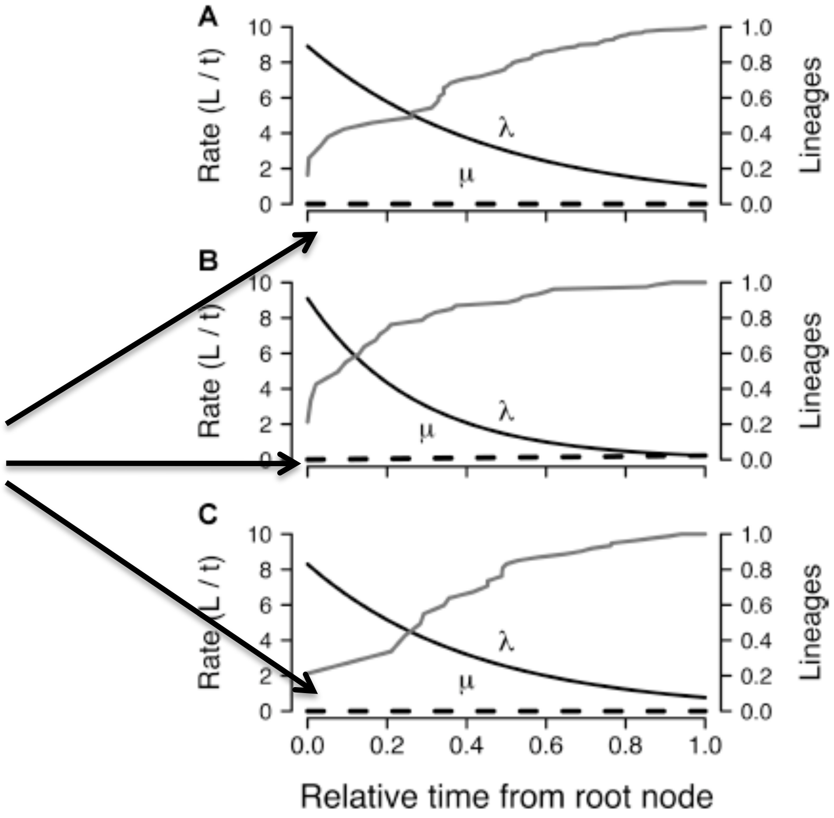


Outline



- My interest in studying long-term diversity dynamics
- How to study long-term diversity dynamics?
- **Current phylogenetic inference is typically inconsistent with the fossil record**
- Reconciling molecular phylogenies with the fossil record
- Conclusions
- Current related project & future directions

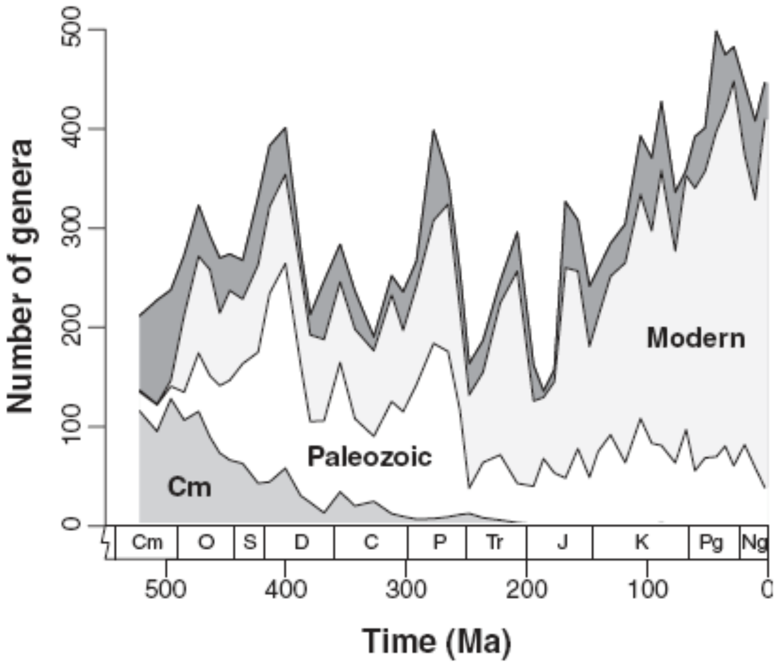
Extinction rates estimated from phylogenies are typically too low to be realistic



Extinction rate

Rabosky & Lovette *Evolution* 2008

Many extant clades should be in decline, but phylogenetic inference does not detect this

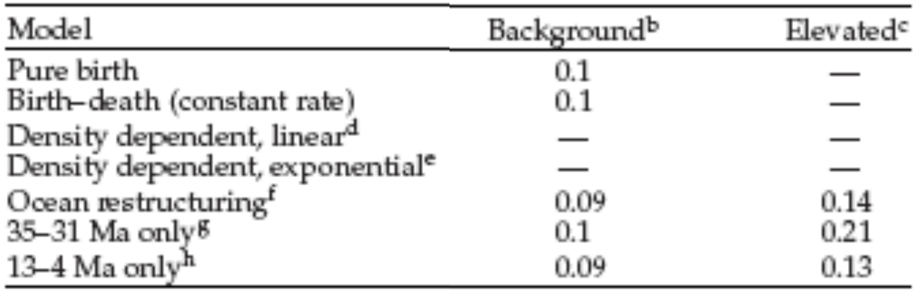


Alroy *Science* 2010

Direct comparison of phylogenetic inference with fossil data show inconsistencies

phylogenetic inference

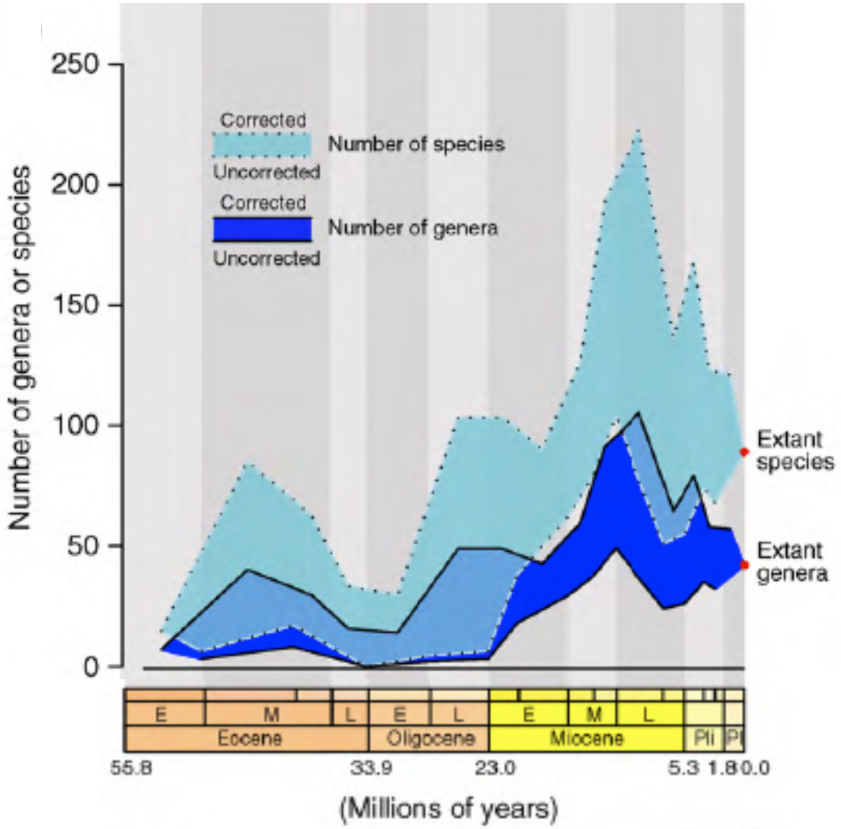
fossil record



Model	Background[b]	Elevated[c]
Pure birth	0.1	—
Birth–death (constant rate)	0.1	—
Density dependent, linear[d]	—	—
Density dependent, exponential[e]	—	—
Ocean restructuring[f]	0.09	0.14
35–31 Ma only[g]	0.1	0.21
13–4 Ma only[h]	0.09	0.13

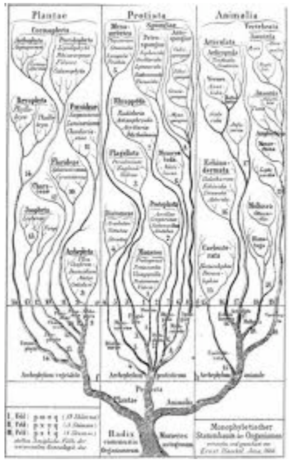
net diversification rates

Steeman et al.
Syst. Biol. 2010



Quental & Marshall TREE 2010

Outline



- My interest in studying long-term diversity dynamics
- How to study long-term diversity dynamics?
- Current phylogenetic inference is typically inconsistent with the fossil record
- **Reconciling molecular phylogenies with the fossil record**
- Conclusions
- Current related project & future directions

Exact likelihood expression for a partially sampled phylogeny under a scenario where speciation and extinction rates are a function of time, conditioned on non-extinction

$$f^{n+1}\prod_{i=1}^{n}\lambda(t_i)\Psi(s_{i,l},t_i)\Psi(s_{i,r},t_i).$$

probability that a lineage survives from t to s and lives no descendant in the sample

$$\Psi(s,t)= e^{\int_s^t \lambda(u)-\mu(u)\,du}\left|1+\frac{\int_s^t e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)\,d\sigma}\lambda(\tau)\,d\tau}{\frac{1}{f}+\int_0^s e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)\,d\sigma}\lambda(\tau)\,d\tau}\right|^{-2}$$

speciation rate

extinction rate

samping fraction

The net diversification rate can take negative values

Morlon *et al.*, *submitted*

Exact likelihood expression for a partially sampled phylogeny under a scenario where speciation and extinction rates are a function of time, conditioned on non-extinction

probability that a lineage survives from t to s and lives no descendant in the sample

$$f^{n+1}\prod_{i=1}^{n}\lambda(t_i)\Psi(s_{i,l},t_i)\Psi(s_{i,r},t_i).$$

$$\Psi(s,t) = e^{\int_s^t \lambda(u)-\mu(u)\,du}\left|1+\frac{\int_s^t e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)\,d\sigma}\lambda(\tau)\,d\tau}{\frac{1}{f}+\int_0^s e^{\int_0^\tau \lambda(\sigma)-\mu(\sigma)\,d\sigma}\lambda(\tau)\,d\tau}\right|^{-2}$$

speciation rate

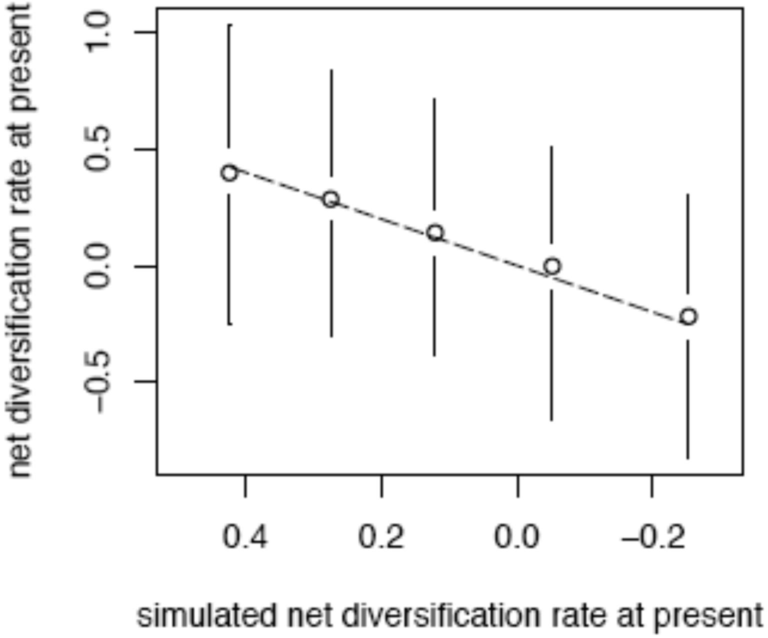
extinction rate

samping fraction

⚠ Speciation and extinction rates are assumed homogeneous across lineages

Morlon *et al., submitted*

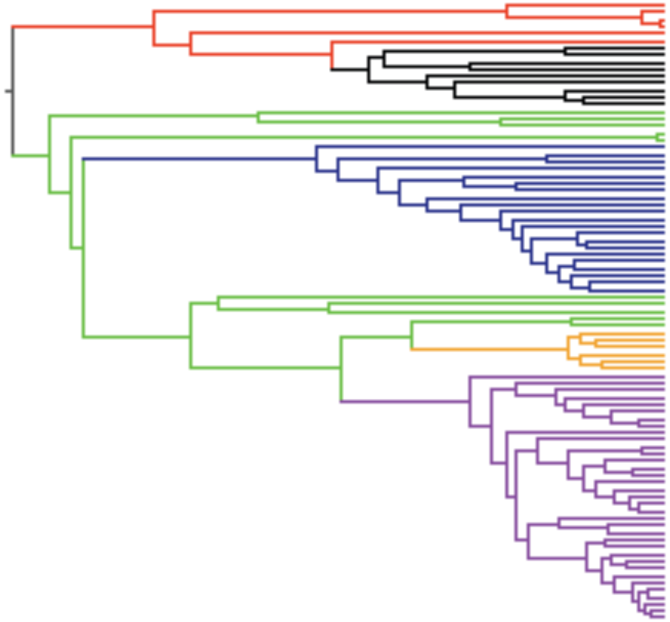
Declining clades can be detected in simulated phylogenies

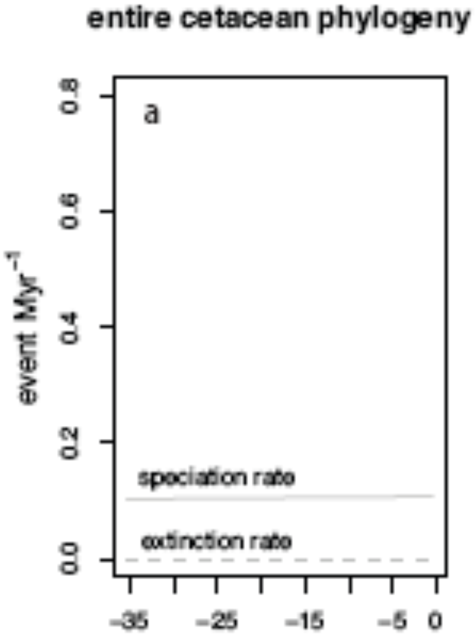


The estimates of speciation and extinction are unbiased

Morlon *et al.*, *submitted*

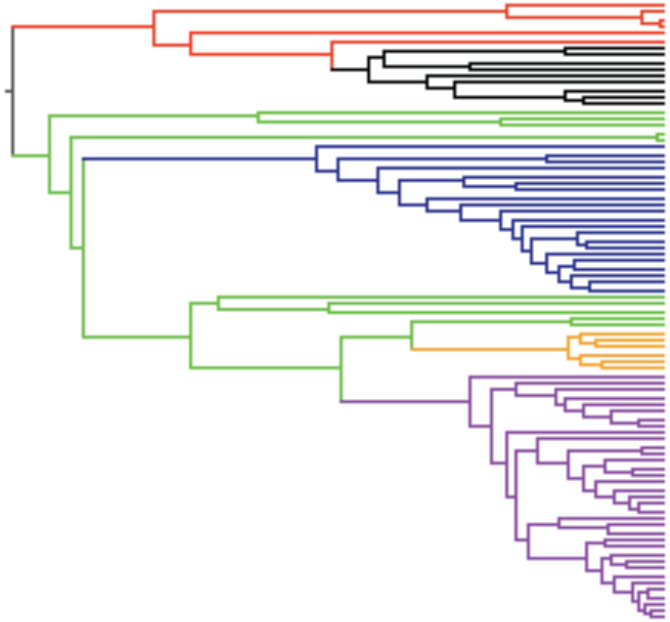
When the cetacean phylogeny is considered as a whole, the diversification rate estimates are inconsistent with the fossil record

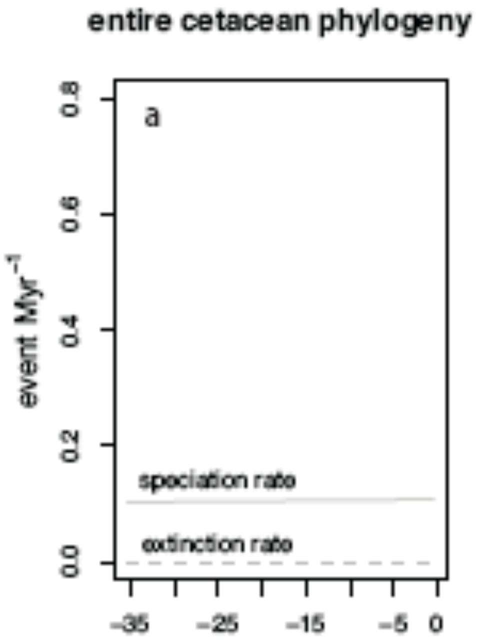




Morlon et al., submitted

When the cetacean phylogeny is considered as a whole, the diversification rate estimates are inconsistent with the fossil record

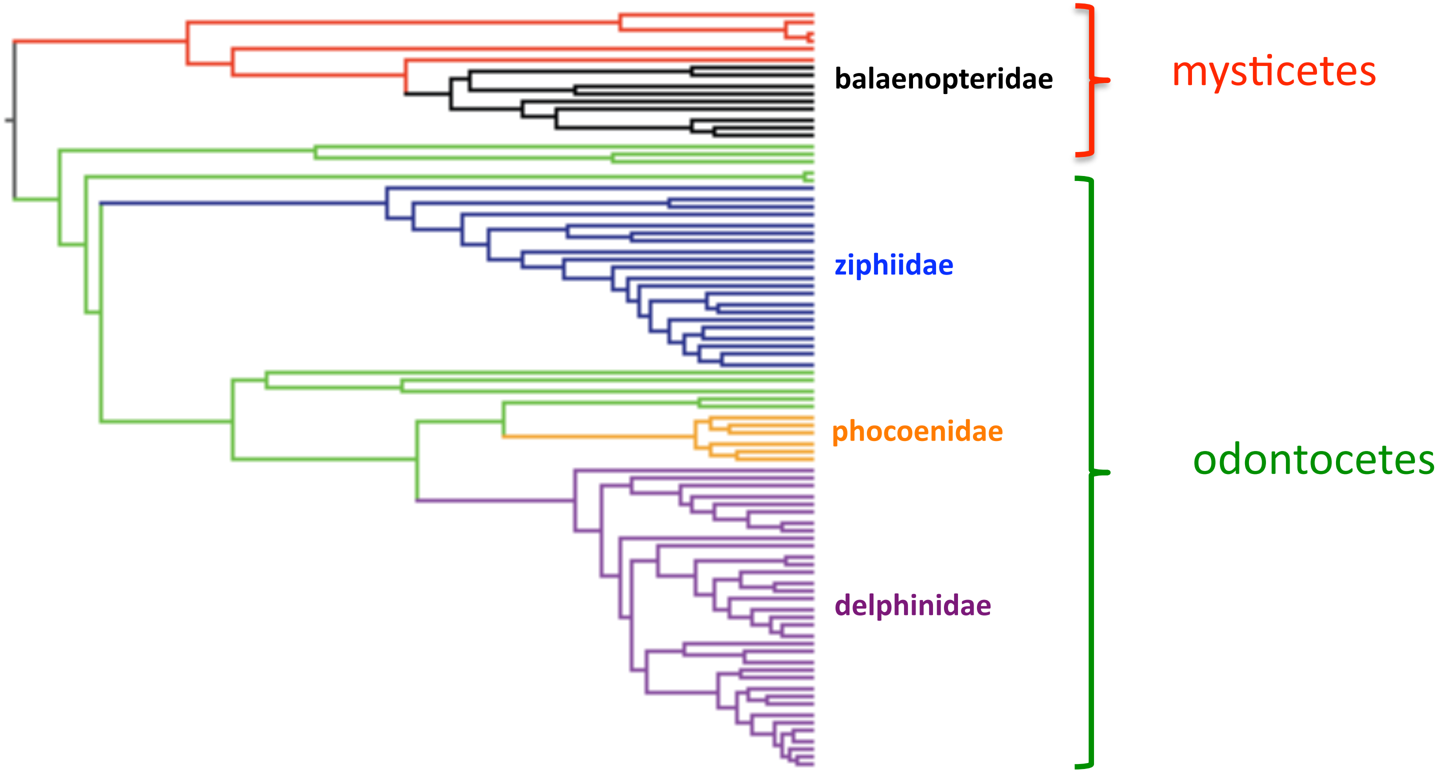




Speciation and extinction rates are assumed homogeneous across lineages

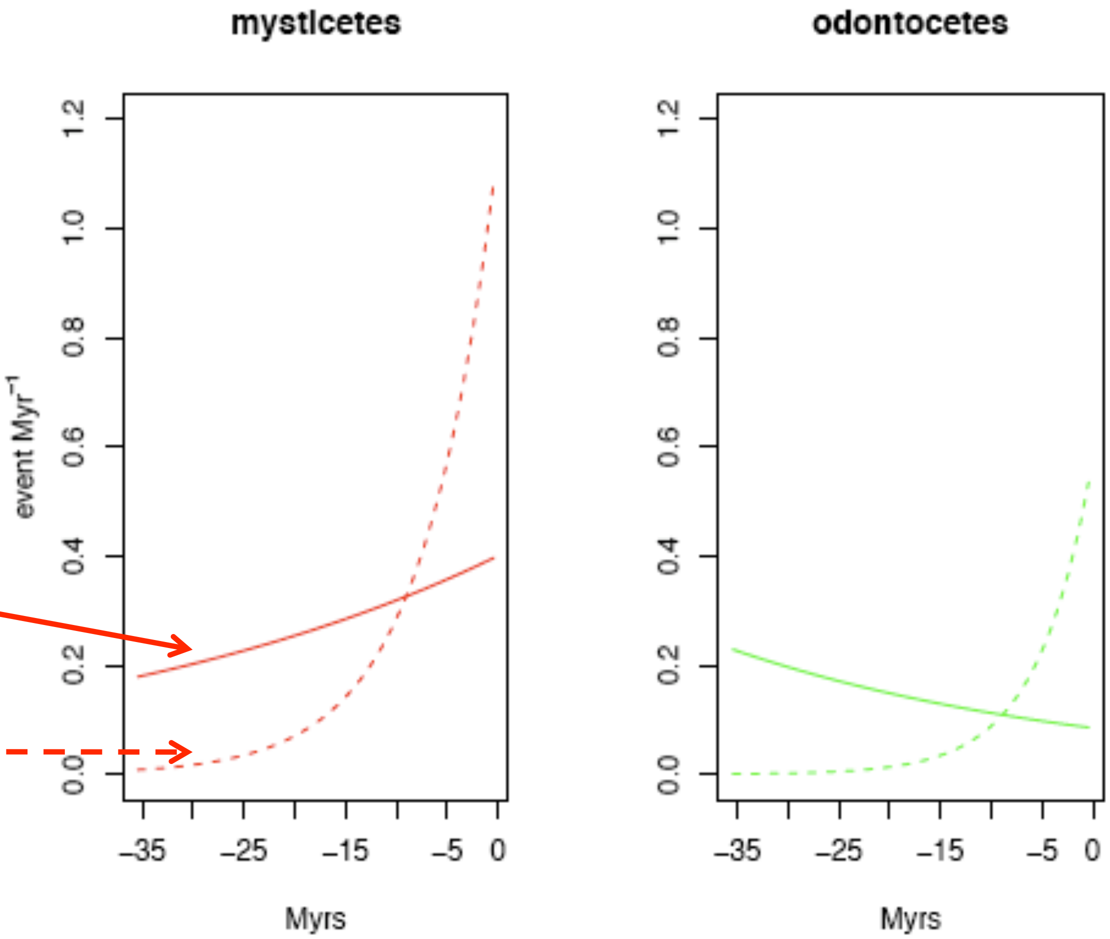
Morlon *et al., submitted*

Out of 14 families, only 4 contain more than 5 species



Morlon *et al.*, *submitted*

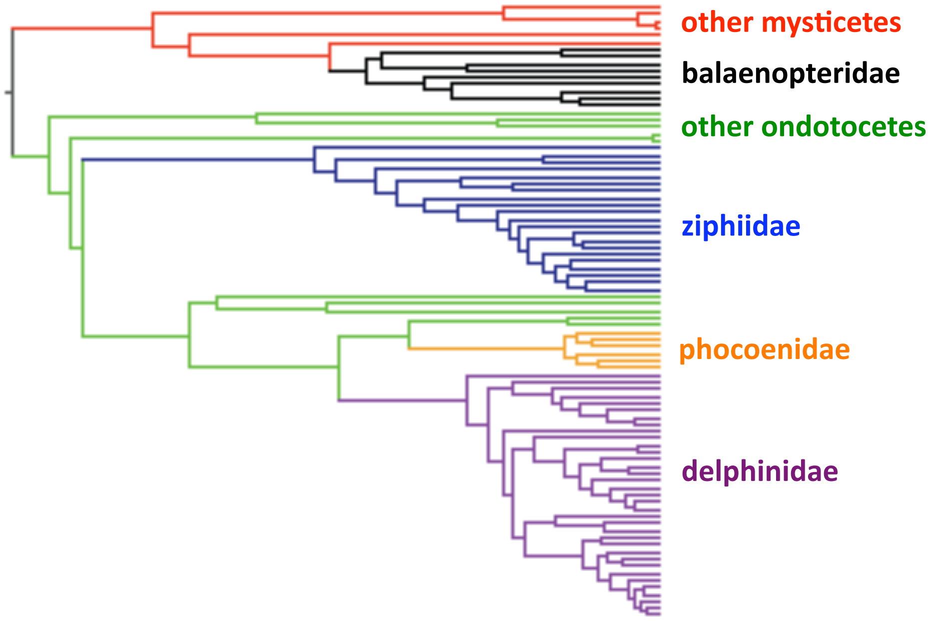
When we isolate the 4 richest families, we infer negative net diversification rates towards the present in the remaining mysticetes and odontocetes

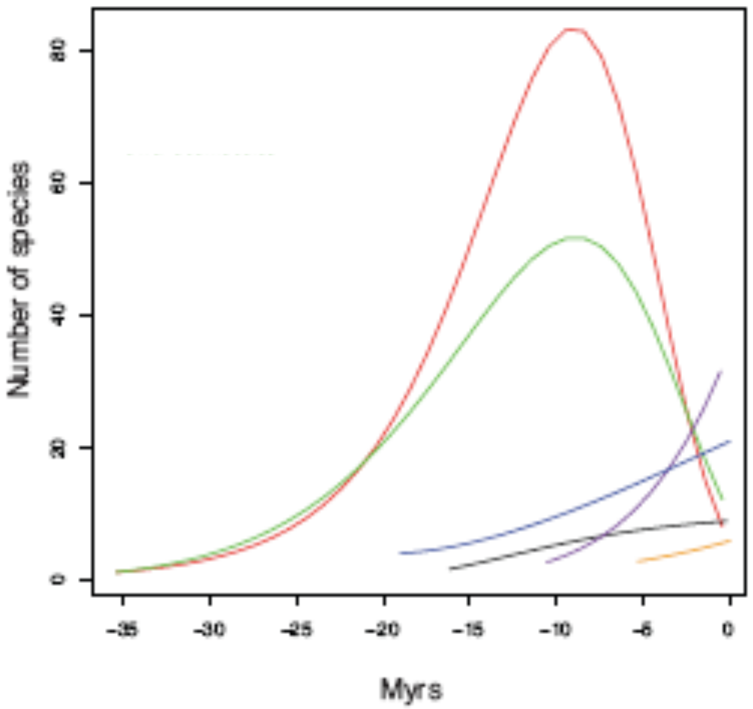


speciation rate

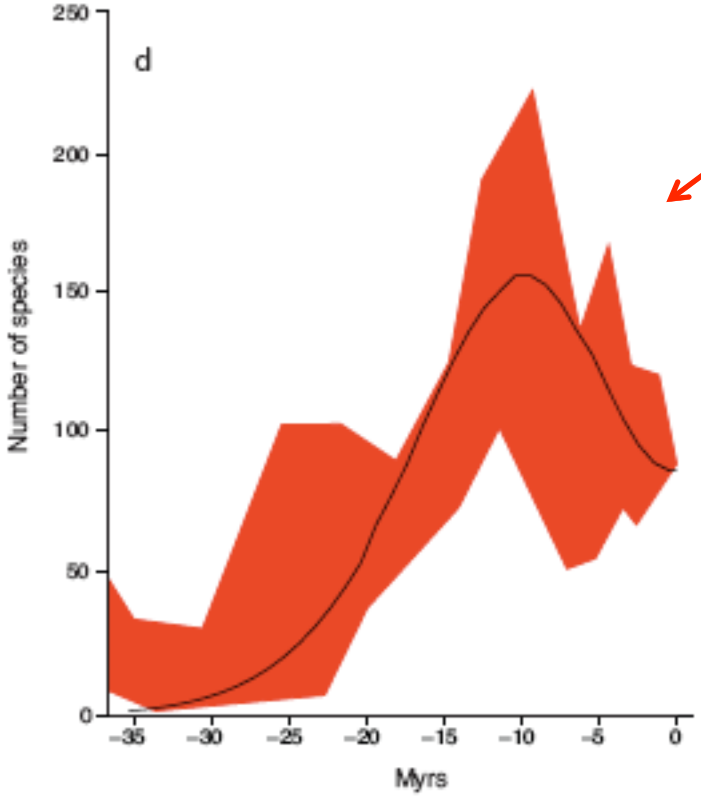
exctinction rate

When we isolate the 4 richest families, we infer boom-then-bust patterns of diversity in the remaining mysticetes and odontocetes





The resulting diversity curve is strikingly consistent with the fossil record



fossil record

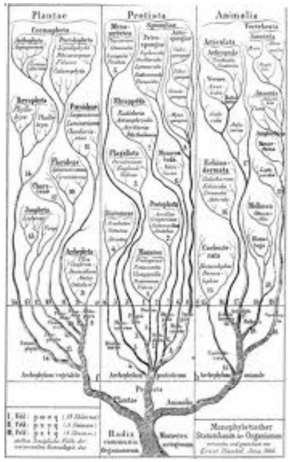
phylogenetic inference, obtained by summing individual diversity curves

Morlon *et al.*, *submitted*

Conclusions

- Our likelihood expression provides unbiased estimates of diversification rates, even for clades in decline
- Realistic extinction rates, and a diversity trajectory strikingly consistent with the fossil record, can be inferred from the cetacean phylogeny, but only if rate heterogeneity is accounted for
- This gives hope for our understanding of diversity dynamics in absence of fossil record
- The biggest challenge is to account for rate variation across lineages

Outline



- My interest in studying long-term diversity dynamics
- How to study long-term diversity dynamics?
- Current phylogenetic inference is typically inconsistent with the fossil record
- Reconciling molecular phylogenies with the fossil record
- Conclusions
- **Current related project & future directions**

Diversity dynamics across latitudes in fabaceae

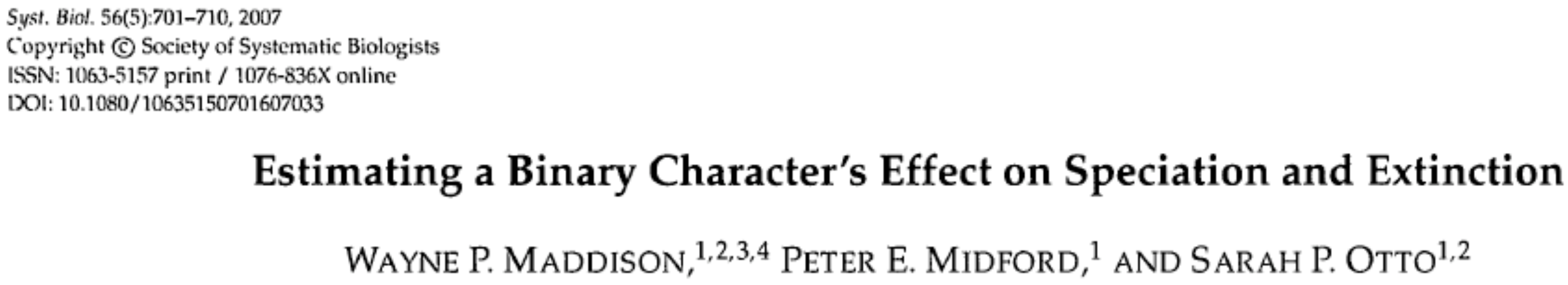
Are speciation rates higher in the tropics? **« cradle »**

Are extinction rates lower in the tropics? **« museum »**

Is the difference in diversity only due to the tropics being older?

Are there less ecological constraints in the tropics?

Method



Syst. Biol. 56(5):701–710, 2007
Copyright © Society of Systematic Biologists
ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150701607033

Estimating a Binary Character's Effect on Speciation and Extinction

Wayne P. Maddison,[1,2,3,4] Peter E. Midford,[1] and Sarah P. Otto[1,2]

Suppose a species can be either in state 0 or in state 1. Each lineage gives rise to a new lineage with rate λ_0 (λ_1), and go extinct with rate μ_0 (μ_1) if it is in state 0 (1, respectively). A lineage in state 0 changes to state 1 with rate q_{01}, and a lineage in state 1 changes to state 0 with rate q_{10}

There is no analytical solution for the likelihood of observing the data (phylogenetic branch-length and character state of extant species), but it is possible to obtain the likelihood numerically, by integration of a series of ODEs

Diversity dynamics across latitudes in fabaceae

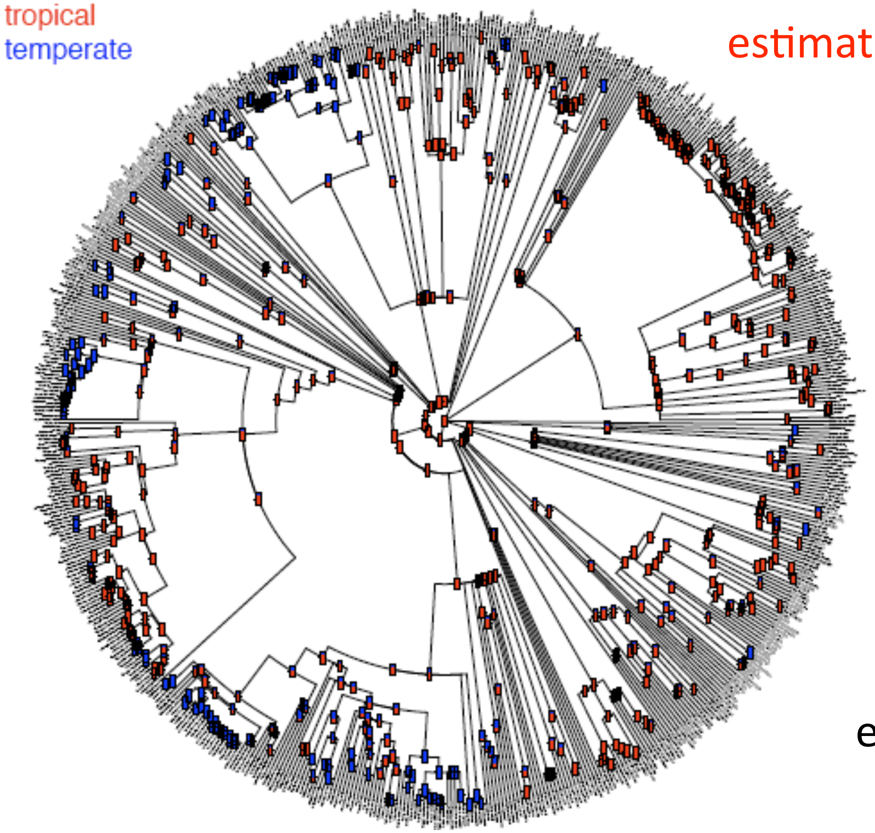
Spatial distribution of fabaceae species across the americas

Brian Enquist, U Arizona

Large phylogeny of fabaceae

Michael Sanderson
U Arizona

Diversity dynamics across latitudes in fabaceae



estimated speciation rate in the tropics (event yr^{-1}): 5.91

estimated speciation rate in the temperate: 5.82

estimated extinction rate: 5.84

estimated rate of transition between biomes: 0.06

Future directions

… as far as theory is concerned

- Accounting for rate variation across lineages
- Accounting for non-random sampling
- Modeling non-Markovian processes (e.g. the rates of speciation and extinction depend on the age of the lineage)
- Trait-dependent speciation/extinction: can we obtain analytical likelihoods? what if speciation and character change are not independent?
- Modeling co-evolution: are diversification rates in one clade dependent or independent of diversification rates in another clade?
- From processes at the individual level to processes at the level of lineages: the neutral theory of biodiversity

THANK YOU!

Joshua Plotkin

Todd Parsons

Dan Rabosky

Charles Marshall

Tiago Quental

Luke Harmon